Generation and validation of spatial distribution of hourly wind speed time-series using machine-learning

Fabio Veronesi[1*], Stefano Grassi[1]
[1]Institute of Cartography and Geoinformation - ETH Zürich; Stefano-Franscini-Platz 5, 8093 Zurich

[*] Corresponding author: fveronesi@ethz.ch

Introduction

Wind resource assessment is a key aspect of wind farm planning in order to assess the long-term electricity production. Moreover, wind speed time-series at fine resolution (i.e. hourly or sub-hourly) are helpful to estimate the temporal change of the electricity generation and indispensable to design stand-alone systems that are affected by the mismatch of supply and demand. In addition, the variability of wind speed is important to evaluate the impact of the electricity injection into the power grid.

The prediction of the hourly mean wind speed time-series is a topic of investigation using both physical and statistical approaches. Physical models take into account factors such as obstacles, pressure, temperature, roughness, and orographic effects. Statistical models correlate wind speed data with remotely sensed, and other physical parameters, to infer the wind spatial pattern. As demonstrated in previous work (Aksoy et al. 2004), statistical methods are computationally efficient, and less time-consuming than physical models, while still generating accurate results. The statistical models include approaches such as ARMA (Castellanos and Ramesar 2006; Philippopoulos and Deligiorgi 2009), Markov chain (Shamshad et al. 2005) autoregressive models (Poggi et al. 2003). The spatio-temporal prediction, i.e. the estimation of the hourly wind speed pattern in areas where no direct observations are available, of wind speed time-series using machine learning techniques is a very recent research topic (Ohashi and Torgo 2012). The major problem in wind resource assessment is the large amount of uncertainty involved from measurements malfunctioning to the extrapolation of the wind speed profile in complex terrains. Only statistical wind resource assessment can precisely account for all these uncertainties.

In this work, we present a new generalised statistical methodology to generate the spatial distribution of wind speed time-series, using Switzerland as a case study. This research is a continuation of the work we presented at EWEA 2015 (Veronesi et al., 2015), and demonstrates that statistical wind resource assessment can be used also for estimating wind speed time-series. In fact, this method is able to efficiently obtain reliable wind speed estimates and to propagate all the sources of uncertainty, from the measurements to the mapping process, so that the final confidence intervals allow a reliable estimation of the stochasticity of the wind resource for a particular site.

Approach

We collected 10 min average wind speed data over 5 years from 161 stations of the network of the Federal Office of Meteorology and Climatology of Switzerland. The wind speed measurements were correlated with around 8’000 predictors, consisting of environmental and climate data with different time and spatial resolutions covering Switzerland (Veronesi et al. 2015).

To account for the anemometer error, we used a resampling approach called Bootstrapping (Efron and Tibshirani 1994) to fit multiple statistical models to the 30 hourly observations (i.e. 6 samples/hour over 5 years). This way we were able to propagate the error from the anemometer data to the wind speed estimations, while also accounting for all other sources of uncertainty. The full estimation uncertainty is represented by confidence intervals around the mean wind speed time-series.

We employed a machine learning model we developed (Veronesi et al. 2016), to generate mesoscale wind speed maps and test its ability in spatially generate reliable hourly wind speed time-series with a robust uncertainty analysis.

Main body of abstract

One of the key issues to plan and successfully efficiently manage wind farms is the ability to know the range of variation of the wind resource in any given time and location. In this work, we used the statistical wind resource assessment approach we developed (Veronesi et al., 2016) to estimate the hourly variation of the wind resource for the entire Switzerland, at 1 km of spatial resolution. We provide practitioners the mean wind speed and the value of the following time-series percentiles: 5th, 10th, 25th, 50th, 75th, 90th, 95th. This allows assessing the range of variation of the wind speed in any given location and time.

The validation of the estimates was carried out with a five-folds spatio-temporal cross-validation. We created a cross-validation loop consisting of 1'000 iterations. For each iteration, 20% of the weather stations were randomly excluded from the training process of the machine learning algorithm. The algorithm was then trained using data of randomly determined subsets of hours of the year. The final step of the loop was the testing phase, where the estimates were compared with the real observations of the excluded locations.

Figure 1 shows the results of the validation process: the monthly bias (i.e. difference between observed and estimated values) for the percentiles 5th and 95th. We present only these values because they provide practitioners with a way to calculate 90% of the hourly variability of the wind resource. On average, this bias resulted -0.08 m/s for the 5th percentile and -0.11 m/s for the 95th percentile. By looking at the plot in Figure 1 we can clearly observe that the variation of the bias of the 5th percentile is relatively small, while it is more acute for the 95th, where the bias is above absolute 1 m/s. In essence, on average the bias of both percentiles is very low, and this indicates that we are able to provide planners with a reasonable estimate of the full range of variation (i.e. 90% of the time the wind has speed within this range) of the wind speed.

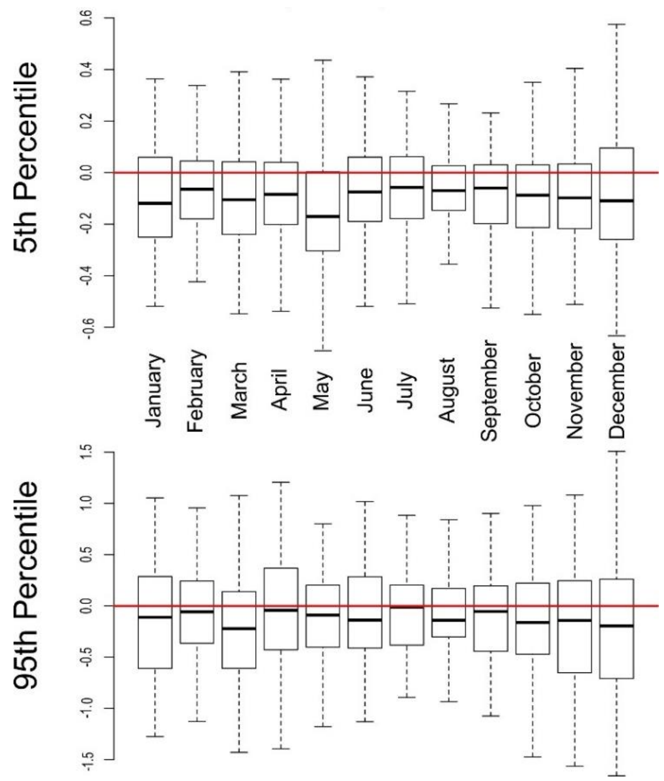

Figure 1: Monthly distribution of the bias of the estimated 5th and 95th percentiles.

Figure 2 shows an excerpt of the generated time series with the mean wind speed (black line) and the percentile 25th, 50th, 75th and 90th (dotted lines). The times-series are used to generate the plot of the temporal variation and the interval of probability of the power output. Figure 3 shows the percentiles of the energy output using a power curve of an Enercon 2.3 MW with a 90 m hub height: the scenarios of the time intervals with missing electricity generation can be identified and used, for instance, to plan the maintenance. The proposed method generates wind speed times series with different probability of occurrence for any selected time window and region useful to investigate to spatial distribution of the impact of potential electricity generation to the power grid and its contribution to the local energy demand.

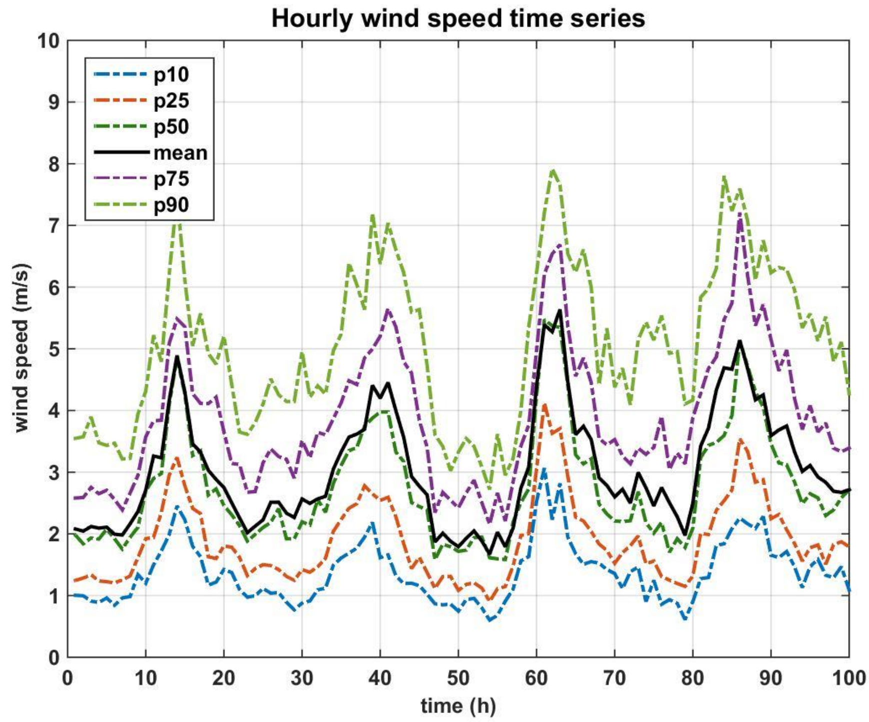

Figure 2: Excerpt of the generated mean wind speed time-series (at 90 m) and the corresponding percentiles.

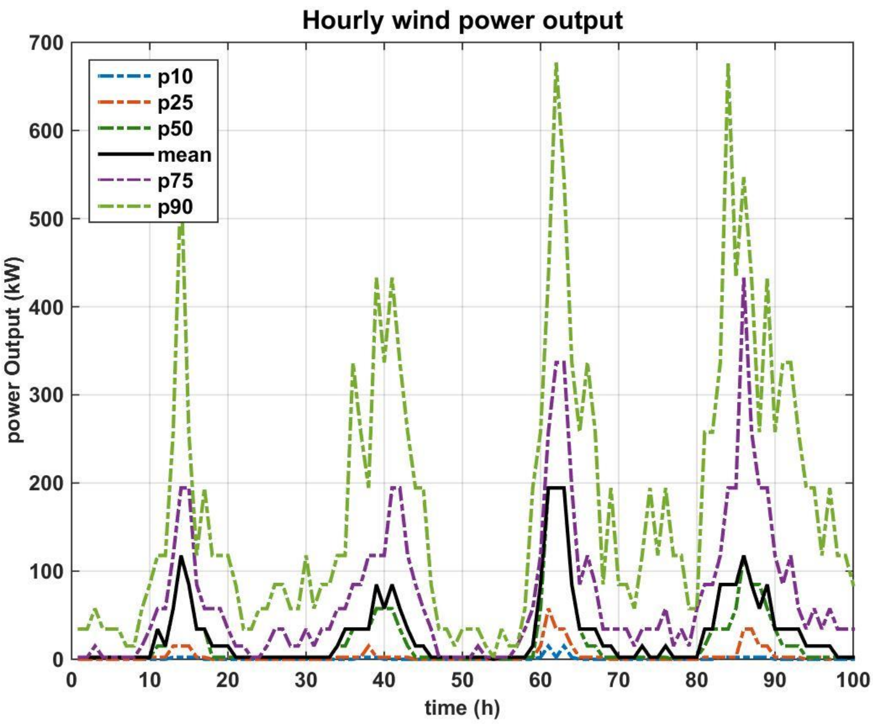

Figure 3: Excerpt of the generated mean power output time-series and the corresponding percentiles.

Conclusions

In this research we demonstrated that the method we developed (Veronesi et al., 2015; Veronesi et al., 2016) can successfully be used to estimate wind speed time-series. Crucially, this method is the only one capable of assessing and propagating every source of uncertainty along the estimation process, to provide practitioners with reliable confidence intervals. Switzerland has been used as case study, but the process can be applied in any other region worldwide.

Learning objectives

This research is an example of the use of statistical wind resource assessment for estimating hourly time-series, including a precise range of uncertainty. Estimating the wind resource at high spatio-temporal resolution using statistical methods is certainly much faster than with traditional methods based on computational fluid dynamics, and this makes this method useful for the industry. In fact, being able to provide practitioners with hourly confidence intervals at high spatial resolution can be advantageous not only for planning but also for wind farms operations.

References

Aksoy, H., Z. F. Toprak, A. Aytek, and N. E. Unal. 2004. "Stochastic generation of hourly mean wind speed data." *Renewable Energy* 29(14):2111-31.

Castellanos, and Vincent Isa Ramesar. 2006. "Characterization and estimation of wind energy resources using autoregressive modelling and probability density functions." *Wind Engineering* 30(1):1-14.

Efron, Bradley, and Robert J Tibshirani. 1994. *An introduction to the bootstrap*: CRC press.

Ohashi, O., and L. Torgo. 2012. "Wind speed forecasting using spatio-temporal indicators." *20th European Conference on Artificial Intelligence (Ecai 2012)* 242:975-80.

Philippopoulos, Kostas, and Despina Deligiorgi. 2009. "Statistical simulation of wind speed in Athens, Greece based on Weibull and ARMA models." *Int. J. Energy Environ* 3(4):151-58.

Poggi, P., M. Muselli, G. Notton, C. Cristofari, and A. Louche. 2003. "Forecasting and simulating wind speed in Corsica by using an autoregressive model." *Energy Conversion and Management* 44(20):3177-96.

Shamshad, A., M. A. Bawadi, W. M. A. W. Hussin, T. A. Majid, and S. A. M. Sanusi. 2005. "First and second order Markov chain models for synthetic generation of wind speed time series." *Energy* 30(5):693-708.

Veronesi, F., S. Grassi, and M. Raubal. 2016. "Statistical learning approach for wind resource assessment." *Renewable & Sustainable Energy Reviews* 56:836-50.

Veronesi, Fabio, Grassi Stefano, and Raubal. Martin. 2015. "Satellite data to improve the accuracy of statistical models for wind resource assessment." in *EWEA*. Paris.